

CELEBAL TECHNOLOGIES PVT. LTD.

ASSIGNMENT

WEEK 7

Let's suppose you have 3 different types of files

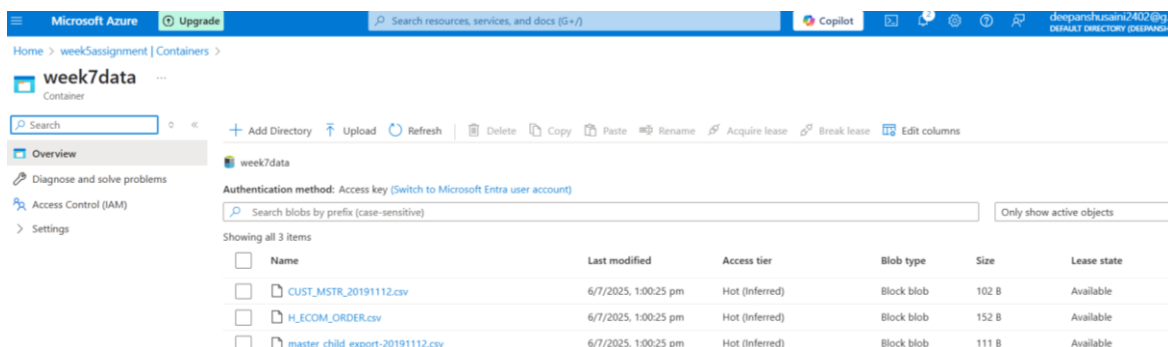
- 1) CUST_MSTR_20191112.csv
- 2) master_child_export-20191112.csv
- 3) H_ECOM_ORDER.csv

All these files will be in the data lake container You have to fetch all three types of files into their respective folders.

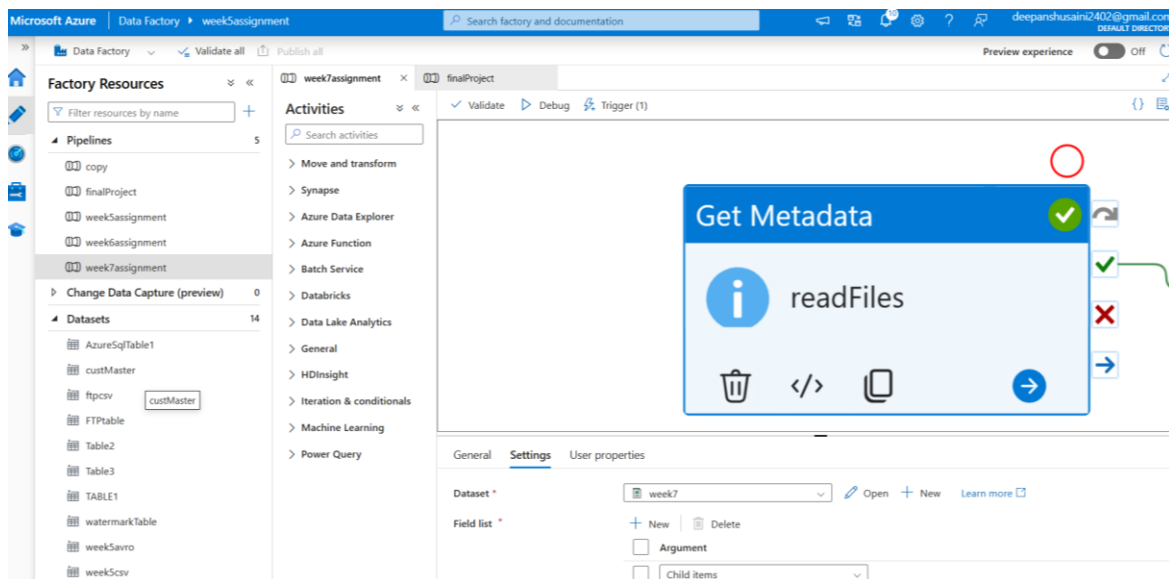
Note: There could be multiple files on all 3 types for different dates for example CUST_MSTR_20191112.csv and CUST_MSTR_20191113.csv

Note: This process will work on truncate load on a daily basis

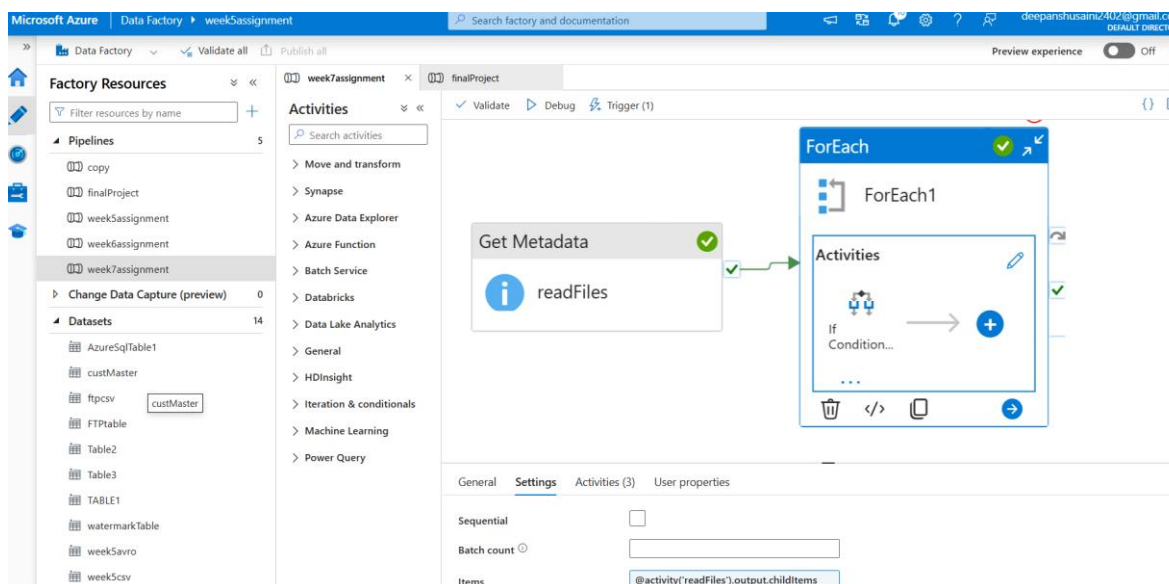
STEP 1: Upload the files in the container in your storage account.



STEP 2: In Azure Data Factory, create a pipeline and add Get Metadata Activity to read files.




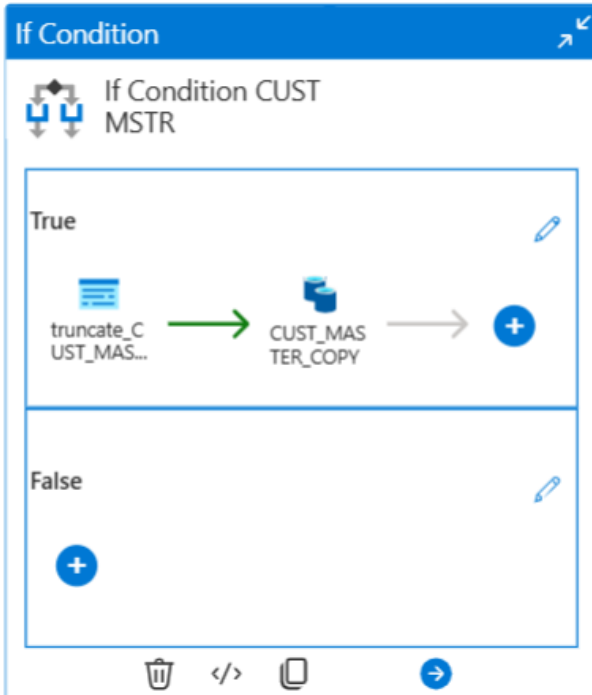
STEP 3: Add ForEach Activity to loop through all files and connect it to Get Metadata Activity.



1) For the "CUST_MSTR" starting name of the file. You have to create an additional column for a date that will fetch the data value from the filename and put it into an additional column Date format: 2019-11-12 and load it into the "CUST_MSTR" table:

STEP 1: Inside ForEach1, add If Condition Activity (If Condition CUST MSTR) with the below given expression.

Preview experience ☒ Off 



Pipeline expression builder

Add dynamic content below using any combination of [expressions](#), [functions](#) and [system variables](#).

```
@startsWith(item().name, 'CUST_MSTR')
```

[Clear contents](#)

ForEach iterator

Activity outputs

Parameters

System variables

...

 Search

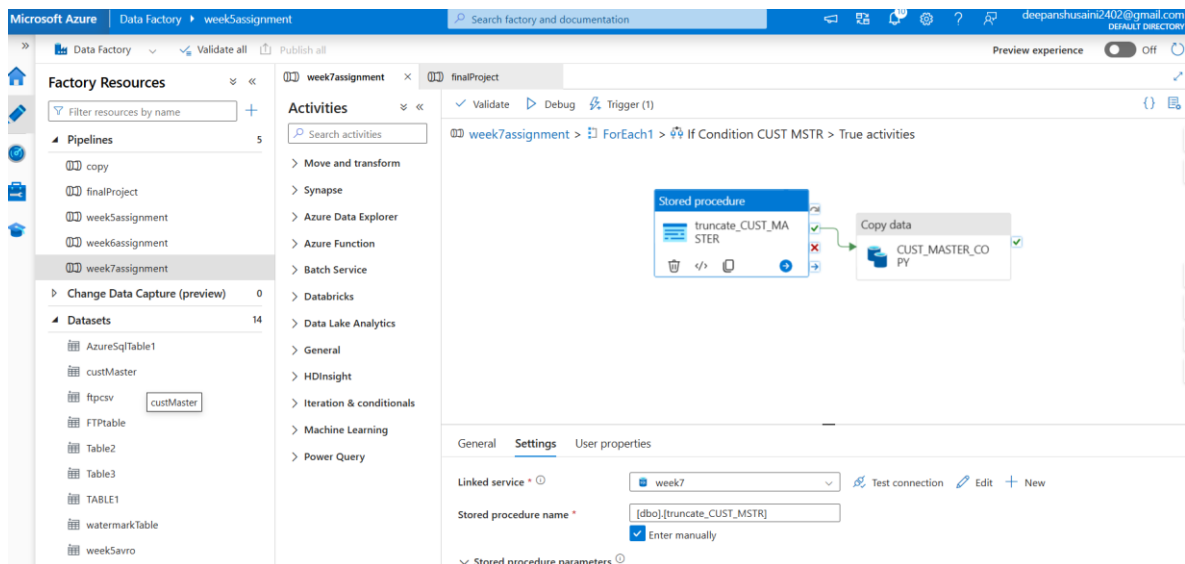
ForEach1

Current item

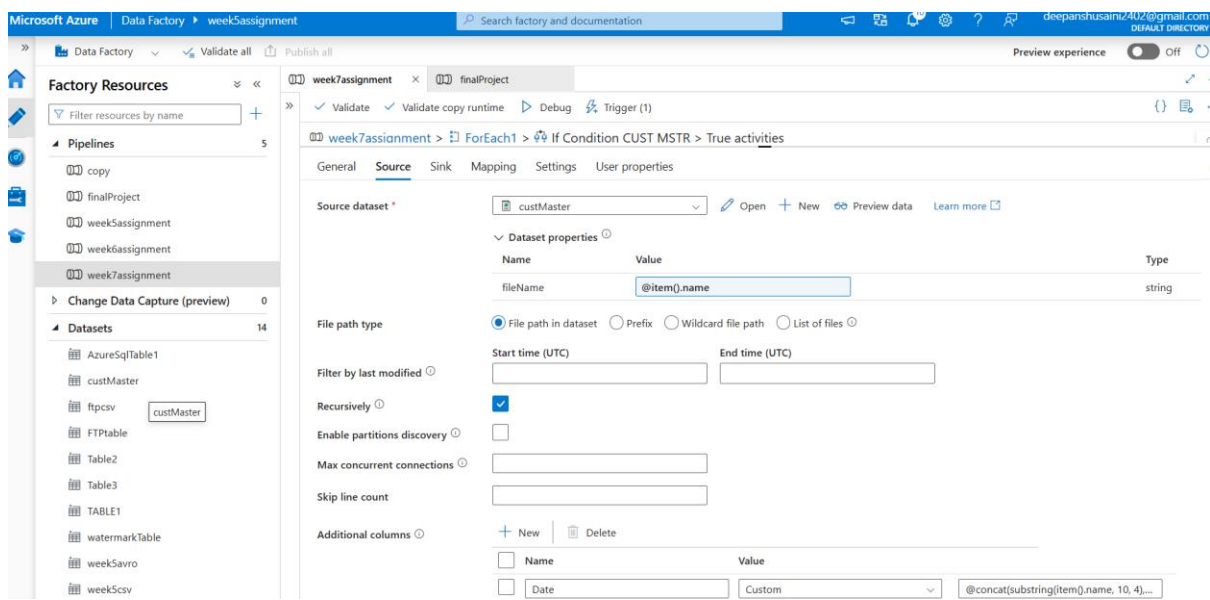
OK

Cancel

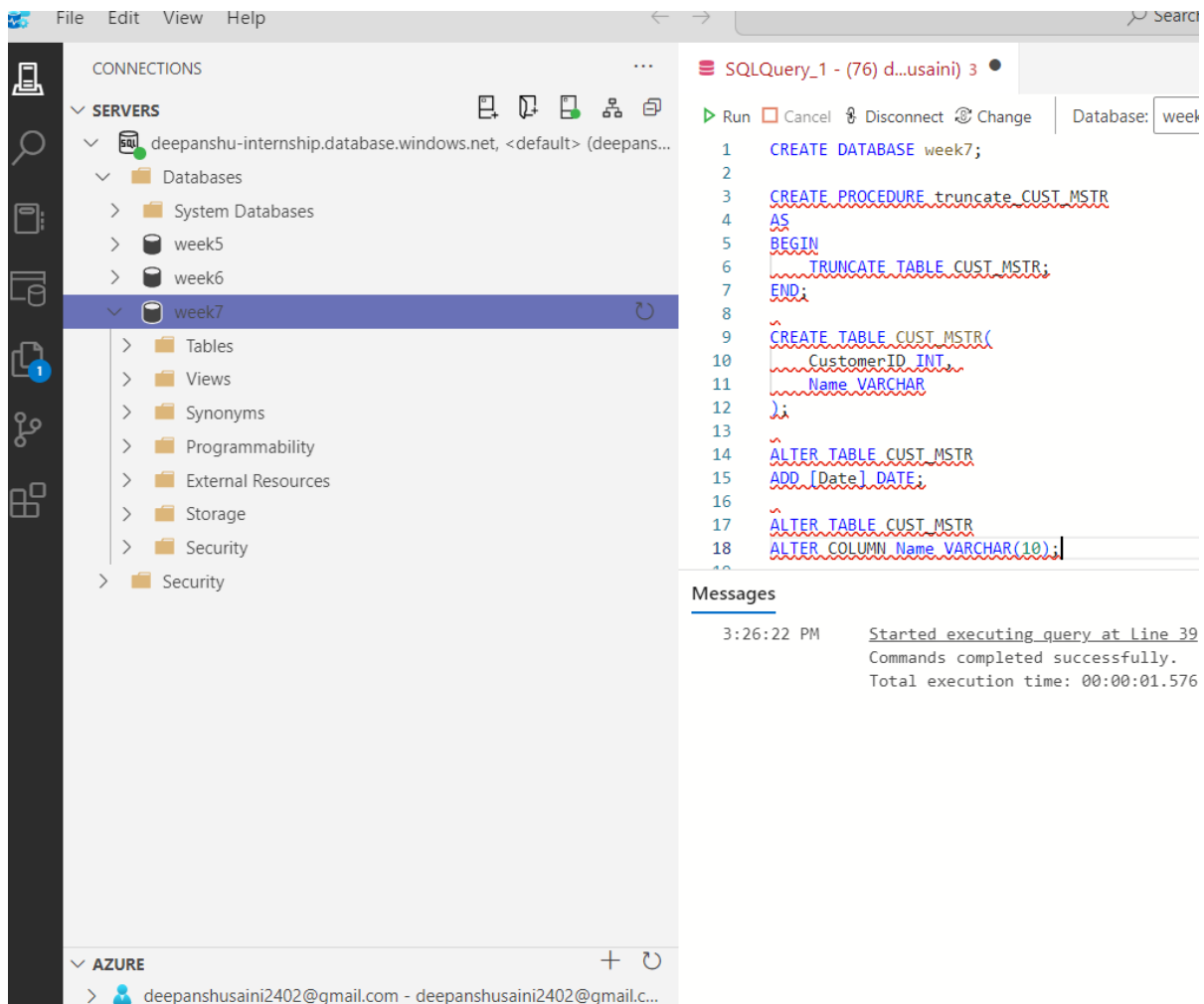
STEP 2: Inside True Case, add Stored Procedure Activity (truncate_CUST_MASTER) and Copy data Activity (CUST_MASTER_COPY), connect the former with latter.



CUST_MASTER_COPY – Source settings



Azure Data Studio Queries:



2) For the "master_child_export" starting name of the file. You have to create two additional columns date and date key which will fetch the data from the filename and put it into the additional columns. Date format: 2019-11-12 DateKey format: 20191112 and load it into the "master_child" table:

STEP 1: Inside ForEach1, add If Condition Activity (If Condition MASTER CHILD) with the below given expression.

deepanshusaini2402@gmail.com

DEFAULT DIRECTORY

Preview experience



Off



If Condition



If Condition
MASTER CHILD

True



truncate-m
aster-child



...

False



Pipeline expression builder

Add dynamic content below using any combination of [expressions](#), [functions](#) and [system variables](#).

```
@startsWith(item().name, 'master_child_export')
```

[Clear contents](#)

ForEach iterator

Activity outputs

Parameters

System variables

...

Search

ForEach1

Current item

OK

Cancel

STEP 2: Inside True Case, add Stored Procedure Activity (truncate-master-child) and Copy data Activity (CopyMasterChild), connect the former with latter.

Search factory and documentation

deepanshusaini2402@gmail.com

Preview experience Off

week7assignment • finalProject

Validate Debug Trigger (1)

week7assignment > ForEach1 > If Condition MASTER CHILD > True activities

Stored procedure

truncate-master-child

Copy data

CopyMasterChild

General Settings User properties

Linked service * week7 Test connection Edit New

Stored procedure name * [dbo].[truncate_master_child] Enter manually

Stored procedure parameters

Import New

CopyMasterChild– Source settings

Search factory and documentation

deepanshusaini2402@gmail.com

Preview experience Off

week7assignment × finalProject

Validate Validate copy runtime Debug Trigger (1)

General Source Sink Mapping Settings User properties

Source dataset * custMaster Open New Preview data Learn more

Dataset properties

Name	Value	Type
fileName	@item().name	string

File path type ☒ File path in dataset ☐ Prefix ☐ Wildcard file path ☐ List of files

Filter by last modified Start time (UTC) End time (UTC)

Recursively ☒

Enable partitions discovery ☐

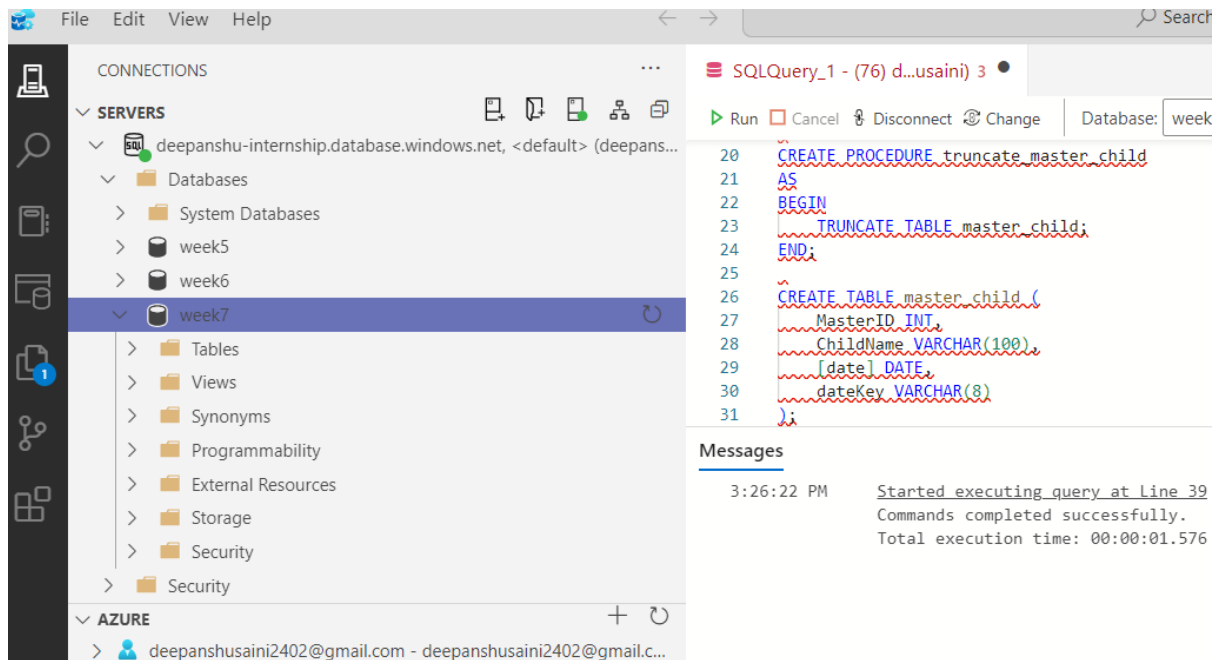
Max concurrent connections

Skip line count

Additional columns

Name	Value
date	Custom @concat(substring(item().name, 21, 4),...
dateKey	Custom @substring(item().name, 21, 8)

Azure Data Studio Queries:



3) for the "H_ECOM_ORDER" type of file you have to load it into the database as it is. and load it into "H_ECOM_Orders" table:

STEP 1: Inside ForEach1, add If Condition Activity (if-h-ecom-orders) with the below given expression.

review experience ☐ Off 



if-h-ecom-orders

True



truncate-h
ecom



...

False



-



Pipeline expression builder

Add dynamic content below using any combination of [expressions](#), [functions](#) and [system variables](#).

```
@startsWith(item().name, 'H_ECOM_ORDER')
```

[Clear contents](#)


ForEach iterator

Activity outputs

Parameters

System variables

...

 Search

ForEach1

Current item

OK

Cancel

STEP 2: Inside True Case, add Stored Procedure Activity (truncate-h-ecom) and Copy data Activity (Copy-h-ecom), connect the former with latter.

Search factory and documentation

deepanshusaini2402@gmail.com
DEFAULT DIRECTORY

sh all

Preview experience Off

week7Assignment • finalProject

✓ Validate ▶ Debug ⚡ Trigger (1)

week7Assignment > ForEach1 > if-h-ecom-orders > True activities

Stored procedure

truncate-h-ecom

Copy data

Copy-h-ecom

General Settings User properties

Linked service * week7 Test connection Edit + New

Stored procedure name * [dbo].[truncate_H_ECOM_Orders] Enter manually

Stored procedure parameters

Import + New

Copy-h-ecom – Source settings

Search factory and documentation

deepanshusaini2402@gmail.com
DEFAULT DIRECTORY

sh all

Preview experience Off

week7Assignment • finalProject

✓ Validate ✓ Validate copy runtime ▶ Debug ⚡ Trigger (1)

week7Assignment > ForEach1 > if-h-ecom-orders > True activities

Copy-h-ecom

General Source Sink Mapping Settings User properties

Source dataset * custMaster Open + New Preview data Learn more

Dataset properties

Name	Value	Type
fileName	@item().name	string

File path type ☒ File path in dataset ☐ Prefix ☐ Wildcard file path ☐ List of files

Filter by last modified

Start time (UTC) End time (UTC)

Recursively ☒

Enable partitions discovery ☐

Max concurrent connections

Skip line count

Additional columns + New

Azure Data Studio Queries:

File Edit View Help

CONNECTIONS

SERVERS

- deepanshu-internship.database.windows.net, <default> (deepans...
- Databases
 - System Databases
 - week5
 - week6
 - week7**
 - Tables
 - Views
 - Synonyms
 - Programmability
 - External Resources
 - Storage
 - Security

AZURE

- deepanshusaini2402@gmail.com - deepanshusaini2402@gmail.c...

SQLQuery_1 - (76) d...usaini) 3

Run Cancel Disconnect Change Database: week

```
33 CREATE PROCEDURE truncate_H_ECOM_Orders
34 AS
35 BEGIN
36     TRUNCATE TABLE H_ECOM_Orders;
37 END;
38
39 CREATE TABLE H_ECOM_Orders (
40     OrderID INT,
41     ProductName VARCHAR(10)
42 );
```

Messages

3:26:22 PM Started executing query at Line 39
Commands completed successfully.
Total execution time: 00:00:01.576